DOCUMENT RESUME

ED 268 555 CS 209 699

AUTHOR Busch, Katharine Mitchell
TITLE Holistic Writing Evaluation: Patterns of Growth in Literacy of Students in the Second and Fourth Grades.
PUB DATE Dec 85
NOTE 25p.; Paper presented at the Annual Meeting of the National Reading Conference (35th, San Diego, CA, December 3-7, 1985).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Elementary Education; Evaluation Criteria; Grade 2; Grade 4; *Holistic Evaluation; Interrater Reliability; Longitudinal Studies; *Writing Evaluation; *Writing Improvement; Writing Processes; *Writing Research; Writing Skills

ABSTRACT

A study explored (1) whether children demonstrated growth in writing ability through the experience of writing two or three times a week without direct teacher instruction; (2) what changes could be observed over time in the writing samples of second and fourth graders; (3) whether written language growth was a continuous progression of upward movement or was sporadic in nature; (4) whether growth in writing would be evident to raters not trained with prototypes of competency levels; and (5) what criteria raters would develop to evaluate growth in writing ability. Samples of writing were taken from ten students as second graders and again two years later when they entered fourth grade. Ten raters were then asked to read the samples and devise their own evaluation criteria. Results showed that the raters were able to identify a pattern of growth in the writing of the second grade students over a 14-week period, but less able to discern growth in the fourth grade writing samples. The results also indicated that the raters had generated criteria in four areas of information: in text, organization of text, language use, and mechanics of writing. Among the conclusions drawn from the study are that growth in literacy can be established within grade levels, that growth and development in writing ability is not the same for all children, and that student writing can improve without any guidance from teachers. (Examples of student writing and raters' responses are analyzed in numerous figures.) (DF)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED268555

# HOLISTIC WRITING EVALUATION: PATTERNS OF GROWTH IN LITERACY OF STUDENTS IN THE SECOND AND FOURTH GRADES

By

KATHARINE MITCHELL BUSCH

California State University
School of Education, PL 532
5500 University Parkway
San Bernardino, CA 92407

Symposium: Assessing growth in student's learning: What non-traditional methods can tell us

Paper presented at 35th Annual Meeting, National Reading Conference, San Diego, CA, December, 1985.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Katharine Mitchell Busch

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."



CS 209 699

Holistic Writing Evaluation: Patterns of Growth in Literacy of Students in the Second and Fourth Grades

Evaluation of writing samples using holistic scoring procedures has gained momentum in several areas of the United States, particulary those areas concerned with competency-based instruction. Holistic scoring as described by Myers (1980) assumes that what is needed first are some prototypes of competency levels, second, some feature analysis of the prototypes themselves, and, third, trained readers who read and score samples of student writing "as a whole or single impression" (p.1). In this procedure, evaluators are given predetermined "anchors" from which levels of competency are based.

One possible alternative to such a system, is to allow for each evaluator to develop his or her own criteria for ranking the text. In this paper, I will discuss the criteria which ten evaluators developed and used for ranking written language of second and fourth graders and the writing profiles which evolved through the ranking process.

The writing samples used in this holistic scoring procedure, were collected in the classroom for a period of two years. The initial phase of the study began with my second grade class and was replicated when these students entered the fourth grade. In each phase, the students were given a series of twenty-four pictures, one at a time, and asked to write a story. The pictures covered a

wide range of topics and settings which took the students fourteen weeks to complete. Given the longitudinal nature of this study, I had the opportunity to record language patterns in written texts within grade levels and across grade levels. I was interested in the following questions:

1. Will children demonstrate growth in writing ability through the experience of writing two or three times a week without direct teacher instruction?
2. What changes over time could be observed in writing samples of second and fourth graders?
3. Was written language growth a continuous progression of upward movement or was the pattern of growth sporadic in nature?
4. Would growth in writing be evident to raters not trained with prototypes of competency levels?
5. What criteria would the raters develop to evaluate growth in writing ability?

To find the answers to these questions, I used a holistic evaluation procedure in which I asked ten raters to read all the texts (a total of 453 samples) written by ten students who had participated both in the second grade phase of the study and the fourth grade phase of the study. The raters were first given all the written texts generated in the second grade, each student individually, and told to sort the texts into two piles. They were required to choose five texts which they believed to be the "first" writing samples and to choose five texts which they believed to be the "last" writing samples. After they had sorted the texts, they

were asked to state the criteria, or traits, they had identified in the written texts which influenced their decisions. As they listed these traits, they were also identifying their own criteria for categorizing texts into sets of "first" and "last."

Second Grade Results

In tabulating the choices of the raters for the second grade students, I developed charts fo each student which listed the raters and the picture number of the first and last five pictures used. This number corresponded with the number assigned to the text. I placed an X under the text number if the rater selected it. If the text number was not one of the original "first" and "last" texts, the space was left blank. I tallied the choices of all ten raters for each student to provide a count of how often the raters were able to choose the "first" and "last" texts. (See Table 5:1)

Table 5:1
Rater Accuracy in Selection of
First and Last Written Texts

Sarah - Second Grade

| | First Texts | | | | | Last Texts | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Raters | 15 | 16 | 17 | 19 | 20 | 8 | 10 | 11 | 13 | 14 |
| 1. H.J. | X | - | - | - | - | - | - | X | X | X |
| 2. K.B. | - | X | X | - | - | X | X | X | X | X |
| 3. K.O. | X | - | - | - | - | X | X | - | X | X |
| 4. B.L. | X | - | - | - | X | X | X | - | - | - |
| 5. R.B. | - | - | X | - | - | X | X | - | - | X |
| 6. W.C. | - | - | - | - | - | - | X | X | - | X |
| 7. R.G. | - | - | - | - | - | - | X | X | X | - |
| 8. S.W. | X | - | - | - | - | - | - | - | X | X |
| 9. D.I. | X | - | - | - | - | X | y | - | X | X |
| 10. S.K. | - | - | - | - | - | X | - | - | X | X |
| Totals | 6 | 1 | 2 | - | 1 | 6 | 7 | 4 | 7 | 8 |

After reviewing the charts I found that, in most cases, raters were unsuccessful in selecting the "first" five texts written by the second graders. However, most raters were succesful in identifying the majority of texts written "last" for seven out of ten second grade students. These results demonstrated that raters were able to observe a strong pattern of growth ovr a fourteen week period for the second grade students.

## Fourth Grade Results

In the case of fourth grade group, most raters were able to identify the "first" text for three out of ten students. While this is a better ratio than the second grade group for "first" texts, raters were not so successful in identifying "last" texts of fourth graders. Most raters were able to identify the majority of texts written "last" for only three out of ten fourth graders.

## Student's Writing Profiles

A second set of charts was constructed which tallied all the choices made by all the raters for each student, regardless of the actual time frame in which the texts were written. These charts were subsequently converted into graphs which revealed the writing profile of the students. In constructing these graphs, I placed the raters' selection of "first" texts below a baseline of zero, which represented those texts not selected by the raters. Through the process of nonselection, these texts represented the group of writings which presumeably were generated in the middle of the

study. Above the baseline were the raters' set of "last" texts. A verical line was placed between the twelfth and thirteen texts which indicated the first and last half of the study. The numbers at the bottom of the graph represented the twenty-four possible texts. Not all of the students wrote twenty-four texts because if they were absent on a particular day, they used the next picture in the sequence. There was no attempt to "make-up" those pictures which were not used because of absence. (See Appendix for writing profiles.)

What was interesting about the writing profiles of the second grade students, was that while raters were unsuccessful in choosing the actual "first" five texts, their choices of "first" texts fell within the range of the first half of the study. Futhermore, when choosing the "last" five texts, they selected texts which had been written in the second half of the study for eight out of ten students. These results were impressive, given the fact the writings were generated over a short period of fourteen weeks in each grade.

In the case of fourth grade students, raters were not as successful in selecting the texts that had been written in the first or second half of the study. Raters were succesful in this task with only three out of ten fourth grade students. The results of this sorting indicated that raters were able to observe a strong pattern of growth in the second grade class, while growth and change appeared to be more sporadic in the fourth grade class. By the term sporadic writing profile, I mean that there was not a

fairly consistent upward movement across the time frame, but that there were peaks and valleys across the series of twenty-four texts.

After reviewing the writing profiles of the students, three patterns seemed to emerge:

1) some students demonstrated a strong writing profile of growth and development in each grade between the first and second half of the study; (See Figure 5:2)

2) some children demonstrated a strong writing profile of growth and development in second grade, but the profile was sporadic in the fourth grade; (See Figure 5:3)

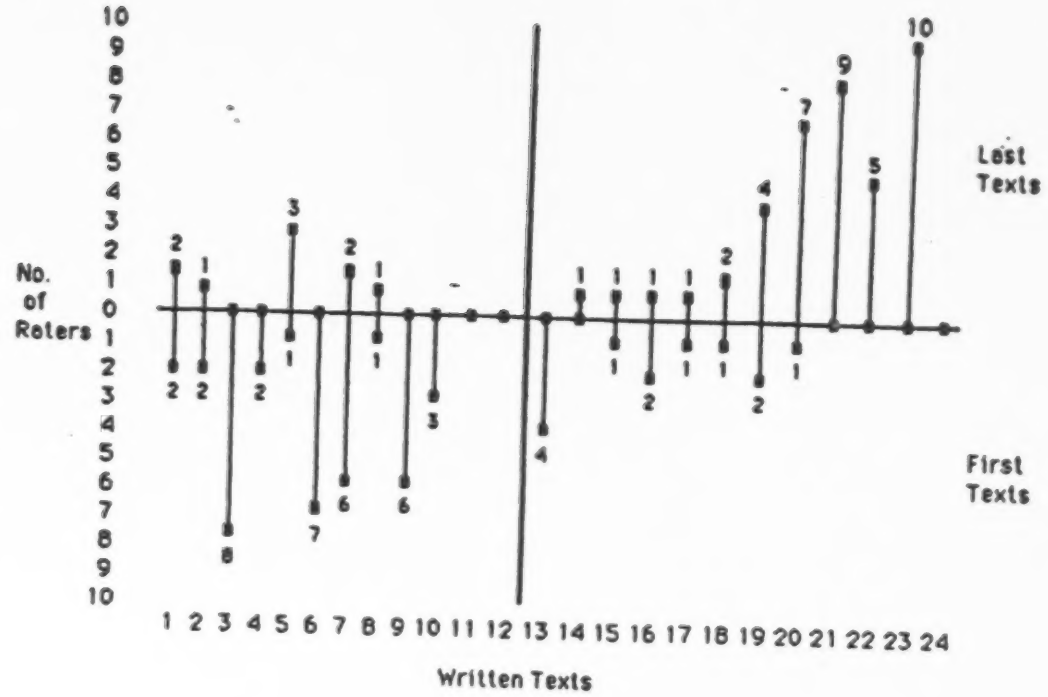
3) some children demonstrated a sporadic writing profile in both grades. (See Figure 5:4)

The writing profile of Christine, Cathy, and Traci illustrate these above patterns.

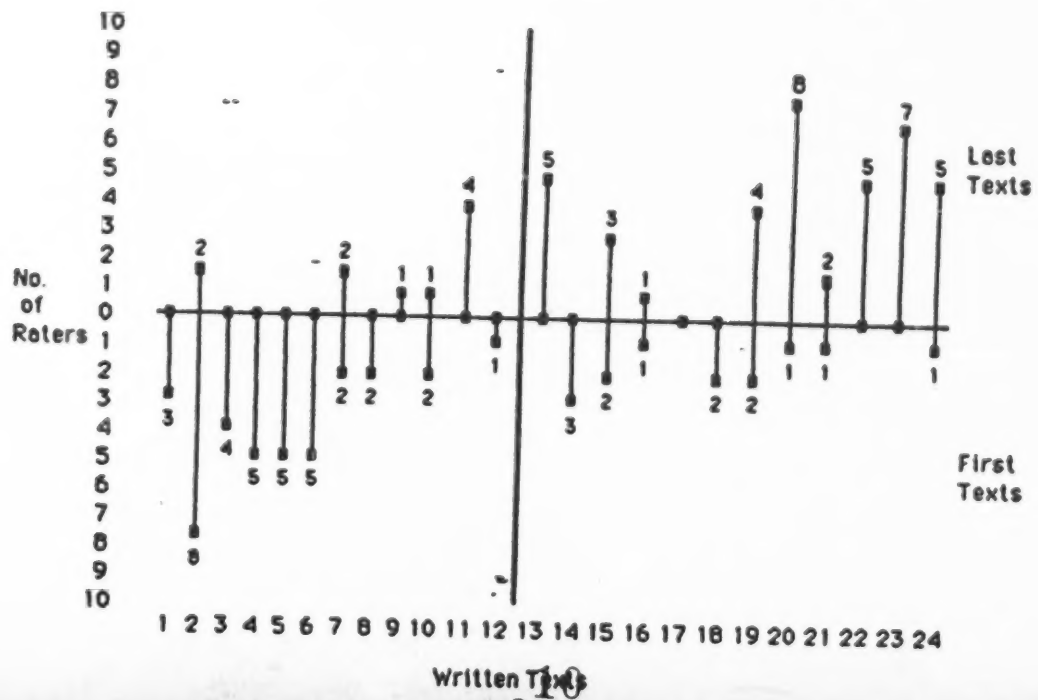
**Rater's Selection of First and Last Five Written Texts**
**14 Weeks**

Figure 5:2

**Christine - Second Grade**



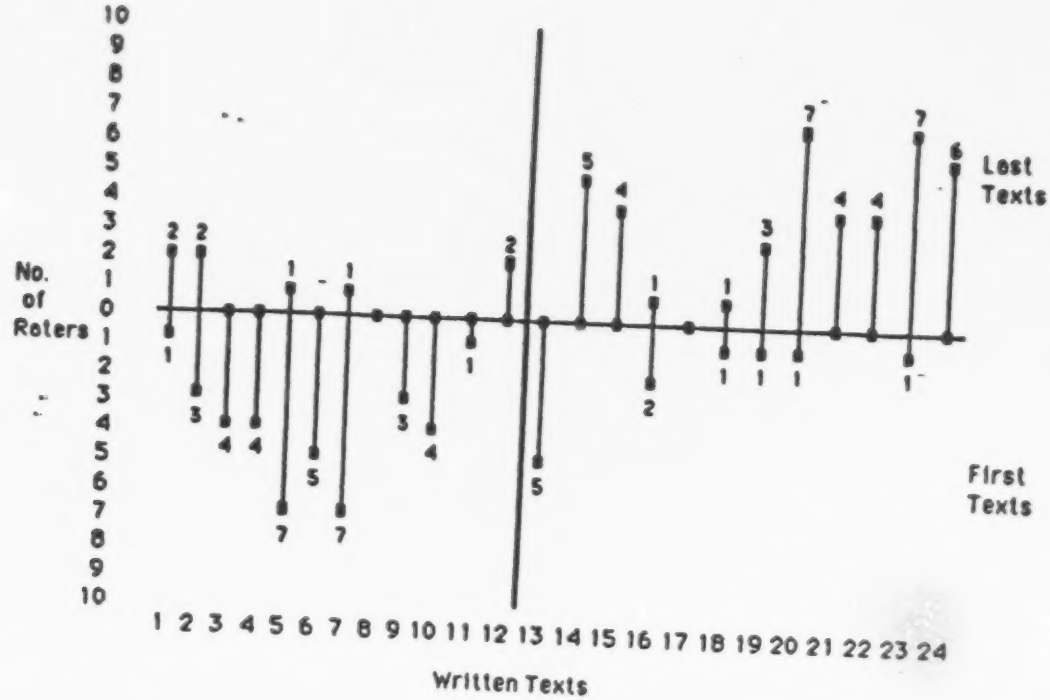

**Christine - Fourth Grade**



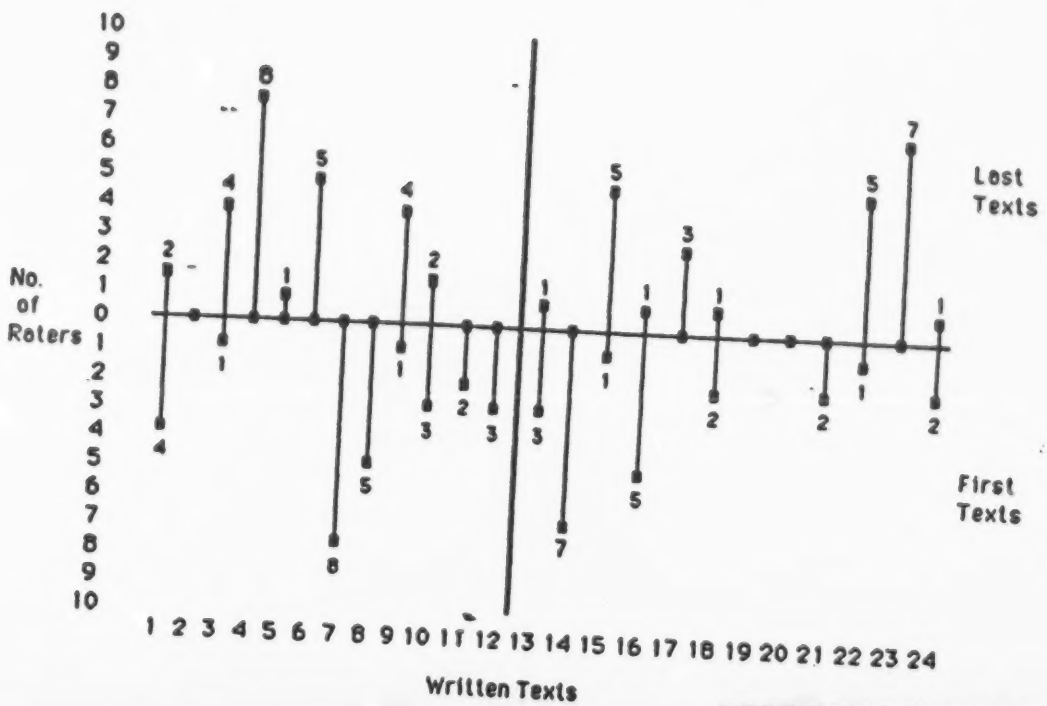
Rater's Selection of First and Last Five Written Texts
14 Weeks

Figure 5:3

Cathy - Second Grade




Cathy - Fourth Grade



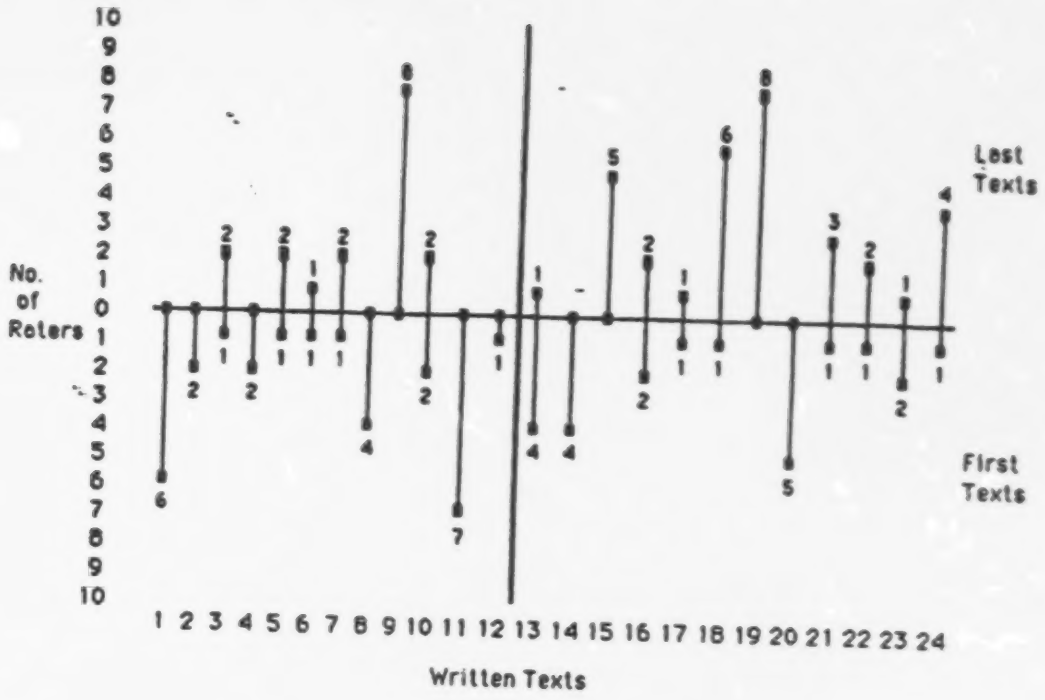



BEST COPY AVAILABLE

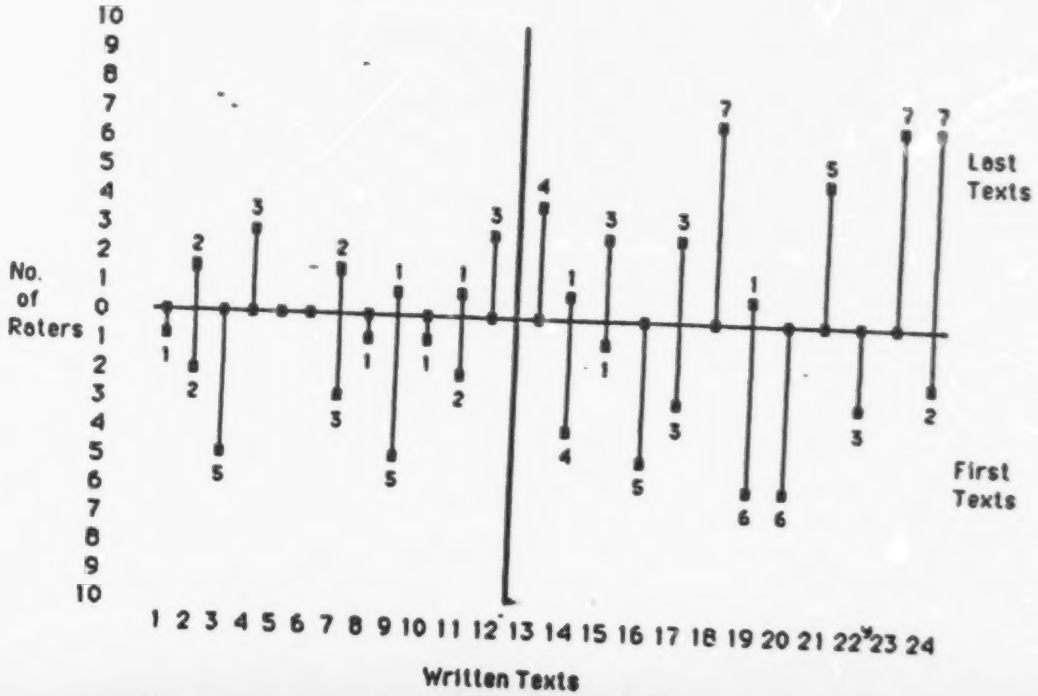
**Rater's Selection of First and Last Five Written Texts**
**14 Weeks**

Figure 5:4

**Traci - Second Grade**




**Traci - Fourth Grade**



In reviewing Christine's writing profile, (See Figure 5:2), while she was in second grade, five or more raters selected four texts written in the first half of the study as "first" texts. Five or more raters were accurate in selecting four out of five "last" texts. The pattern was similar for Christine as a fourth grader. Five or more raters selected four "first" texts which fell within the first half of the study. The same pattern holds true for the "last" texts which fell within the last half of the study. Christine's writing profile demonstrated a pattern of growth both in the second grade and in the fourth grade.

Cathy's writing profile (Figure 5:3), illustrated a different pattern of development. While she was in the second grade, raters identified three texts which were written in the first half of the study and four texts which were written in the second half of the study. While Cathy was in the fourth grade, raters were generally more in agreement as to which texts they considered to be written "first" and "last," but they were less accurate as to the time frame in which the texts were actually generated. So while Cathy's writing profile showed a strong pattern of growth in the second grade, her development was more spordic in nature in the fourth grade.

The writing profile of Traci (See Figure 5:4), was different from the two previous students. In her case, raters were unable to determine the time frame in which her texts were generated both in the second and the fourth grades. In reviewing her fourth grade profile, it appears that most of the raters were in agreement that

three of her texts were generated in the second half of the study. However, they also chose three texts as "first" text that were actually written in the second half of the study.

There may be several reasons that raters were able to discern growth in writing ability more easily in second graders than in the fourth graders:

1) second graders are learning about language and the writing process through engaging in the writing event, so it is easier for observers to see growth in writing ability;

2) fourth graders have more language under their control so a continuous pattern of growth may not be so apparent for them as it was for the second graders;

3) fourth graders may be more confident about their writing ability and therefore be better risk-takers;

4) fourth graders may be more likely to exercise their option not to fully engage in the writing event if they have a picture which does not interest them.

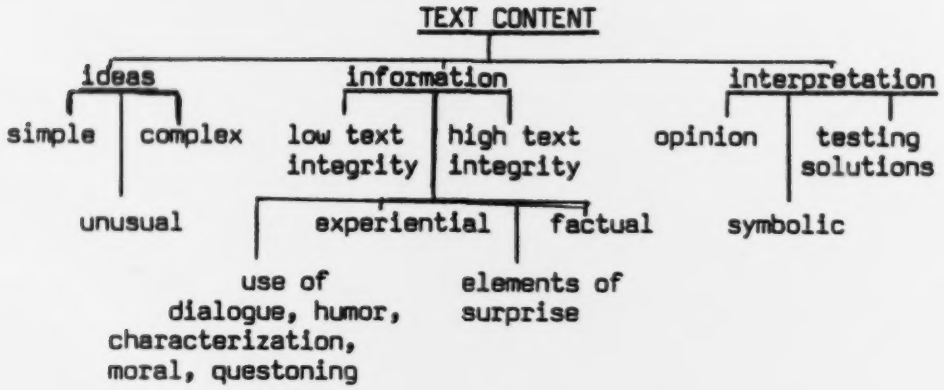
The sporadic writing profile in the second and fourth grade populations is similar to our own progress in writing ability. Elbow (1973) has observed that writing badly is a crucial part of learning to write well. He notes that: regressing and falling apart are a crucial and necessary part of any complex learning." In Graves' study (1983), he found that not every piece of writing was better than before even though there was general improvement in writing ability.

Rater's Criteria for Observing Growth in Writing

In order to compile the rater's criteria for growth in both grades, I went through their lists of criteria and wrote each item on a separate card and continued this procedure until all the items were on cards. I then sorted the cards into categories. I found that raters had generated criteria in four areas: information in text; organization of text; language use; and mechanics of writing. (See Figures 5:5 and 5:6)

Under the category of Text Content, raters listed items regarding the complexity of ideas, elaboration of information; use of imagination, opinions and interpretation of picture text; and high/low text integrity. (See Figure 5:7)

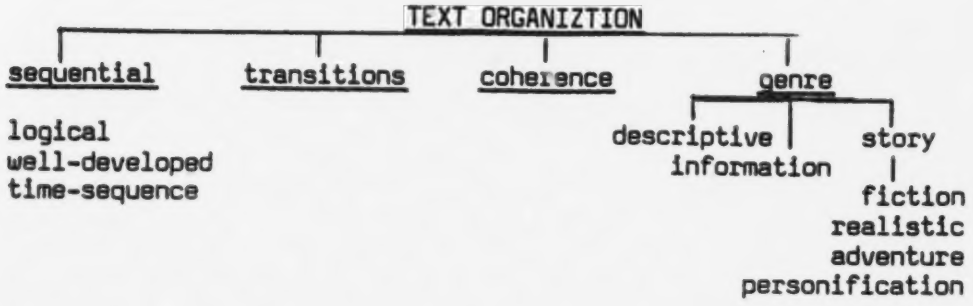
Figure 5:7




Raters' criteria for "first" texts were those with simple ideas, with heavy dependance upon the picture using a listing or retelling of facts with little personal investment. In contrast, raters' criteria for "last" texts were those texts which contained unusual or complex ideas, with little dependance on the picture text, using dialogue, humor, characterization, or testing solutions. For the most part, second grade texts were experiential in nature. However, the "last" texts were viewed to have more personal investment on the part of the writer.

Included in the category of Text Organization, were items regarding the logical structure of the texts; coherence among picture elements; transitions, and genre. (See Figure 5:8)

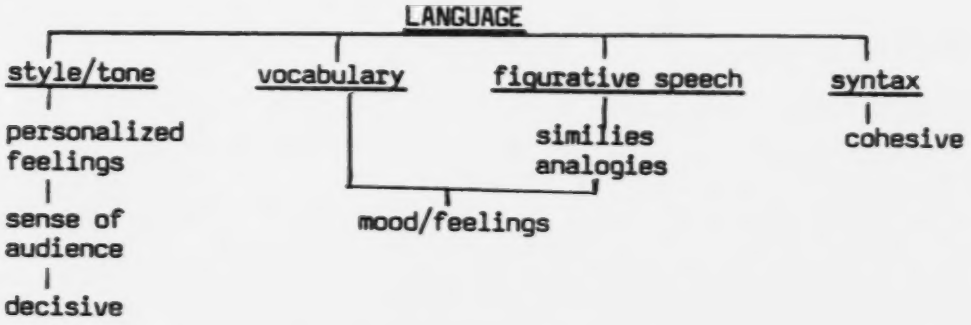
Figure 5:8




Raters' criteria for "first" text centered on the lack of text organization, and the concentration of the genre in the descriptive or informational mode. Passages that were classified as stories by the raters were more apt to be placed in the "last" text category.

Under the category of Language were style and vocabulary, figurative speech, simple or complex syntax; and cohesiveness. (See Figure 5:9)

Figure 5:9




In the area of Language, raters viewed "first" texts as containing simple vocabulary, disjointed sentences, with no sense of audience. Some raters used the term "egocentric" language. By

contrast, texts placed in the "last" set contained a richer vocabulary with the use of similies and analogies. The syntax was judged to be more cohesive and the reader had a sense of "knowing the writer better."

The items included in the category of Mechanics were spelling and text length. Spelling became a criteria for some raters because in the second grade texts the spellings were conventionalized and in the fourth grade texts they were not. This difference was because the second grade texts were typed from audio tapes and the fourth grade texts were typed from original manuscripts. Text length was also included in this category because in my view, that item is related to the mechanics of writing. However, I concede that there is also an arguement that length of text is related to the use of language.

Reviewing the range of criteria generated by the raters in their effort to rank the texts into sets of "first" and "last," I became aware that young writers are more competent in language and language use than we generally assume.

Patterns of Growth Over Time

Although raters were asked to choose the "first" and "last" texts for each student, that request was translated by the raters into the task of ranking high, middle, and low writing samples. The assumption which seemed to underlie the sorting process was that better quality texts would more likely be produced at the end of the study than at the beginning of the study. This assumption

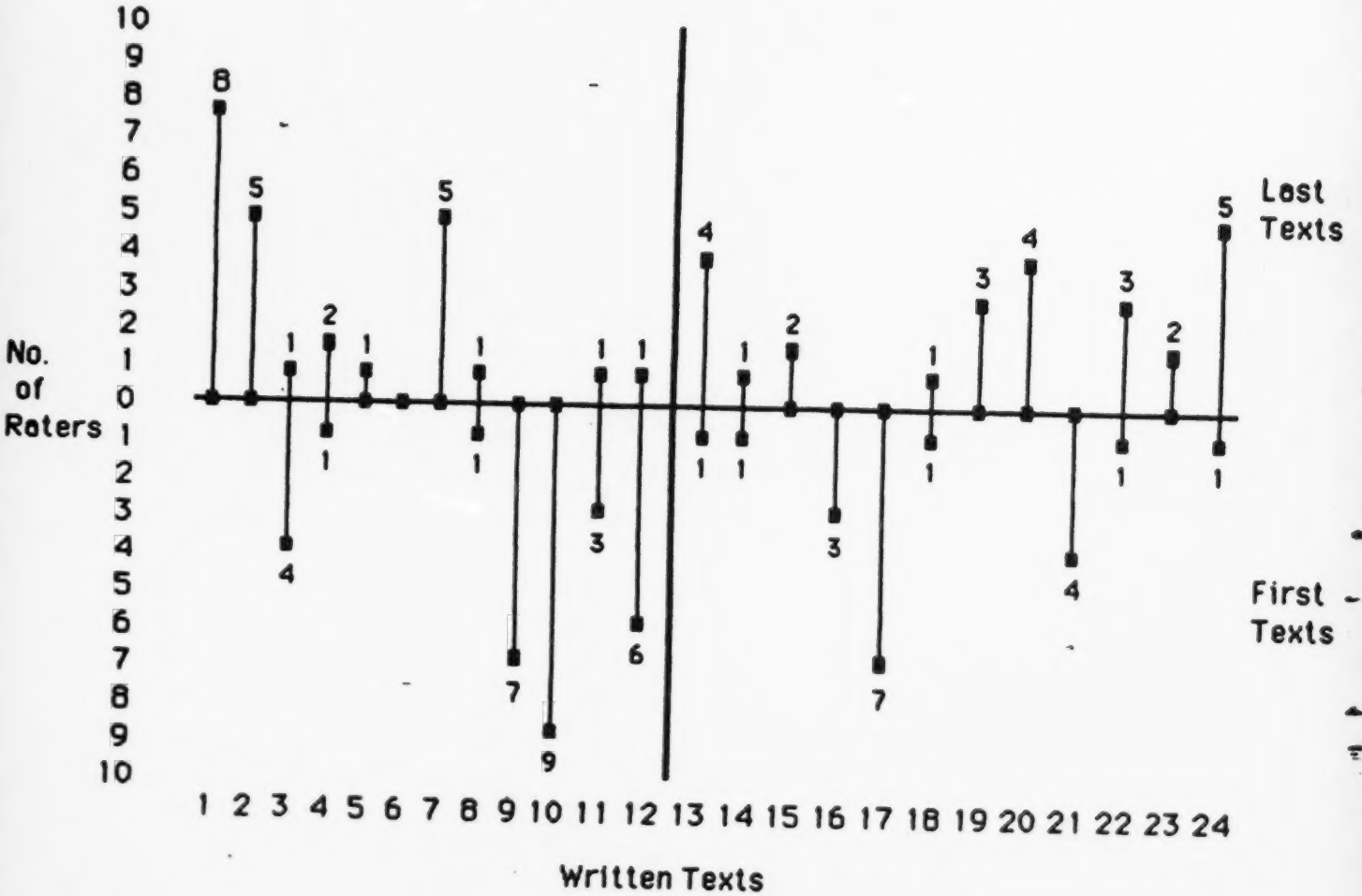
was so strong that even when raters said that shorter or less developed texts were not necessarily written at the beginning of the study, those features would ultimately be the deciding factor in the selection of "first" texts when the choice between two texts became difficult.

This strategy of choosing longer texts and more complex sentences worked fairly well for most second grade texts. However, it soon became apparent to the raters that there was a noticeable difference between second and fourth grade writing samples. After sorting the second grade texts, most raters commented that the fourth grade texts were more difficult to rank as the texts were basically more sophisicated and it was difficult to tell when the texts were generated. These feelings are reflected in the sporadic nature of the writing profiles for the students in fourth grade. Many texts, which were longer and seemed more mature, were actually written early in the study. Also, length and complexity of texts were not the only trait which influenced the raters. The element of interest was a major factor when ranking the fourth grade texts. If the writing "appealed" to the rater, s/he would be more likely to place that piece in the "last" text category. (See Figure 5:10) Several writing samples were ranked as "last" texts, when in reality, they were written in the first half of the study.

Figure 5:10

**Rater's Selection of First and Last Five Written Texts**
**14 Weeks**

**Dierk - Fourth Grade**



All of the raters believed that saw growth in writing ability across grade levels. There were several reasons for this. First, the surface features of the texts were different. As I have noted earlier, the second grade texts were typed from audio tapes made when the second graders read their story to me. Therefore, all the spellings had been conventionalized when those texts were typed. In the case of the fourth grade texts, they were typed directly from the manuscripts with the functional spellings and original punctuation in tact. So while handwriting was not a factor in discriminating between these two grade levels, or within grade levels, spelling and punctuation were a major variable. In addition, the length of texts were noticeably different. Second grade texts were ususally four or five lines in length, while most of the fourth grade texts filled at least a half a page, often more. This is not to imply that fourth graders, or adults for that matter, do not write short texts. It was the combination of these surface level features that led all the raters to state emphatically that they could tell second grade texts from fourth grade texts at a glance.

The writing samples of two students, Eric and Sarah are provided to illustrate the differences in the surface features as well as the content across grades:

Eric - Second Grade Picture #7

Maybe these are cement blocks and someone is going to buy them and build a statue on the stool behind the blocks.

Eric - Fourth Grade Picture #7

The Roman Stackhue

This is an old stache of a Roman knight. The stachue can be seen in Paris. It is in the famous museum along with some of the world's most beautiful and fabulous art work. This stachue was by hand, but who ever made a fabulous piece of art is not know. This stachue is about 250 years old or more. This stachue was made of stone and mud and there isn't anything else in it. Shap pointy rocks were used to carv all of the shapes in it. This big stachue was found in a underground cave. This stachue is sead to be worth 2 billion dollors. There have been avout 50 other people who have tried to by it for about 250,000 dollar but didn't get it.

Sarah - Second Grade Picture #21

This is a pretty picture. It has red and white. It has blue and it has little speckles of blue. And it has little speckles of red. It is pretty colors and it has colors that are mixed together. It has red and blue mixed together. It looks like someone painted it. It has white and red mixed.

Sarah - Fourth Grade Picture #21

Once upon a time there lived twin sisters and a mom but the dad died. He had fell in a pond behind there house. Some people said that one you fall in the pond you will never return they said you will become a slave of the devil. And it was true. One day the twins were going outside to play and one of the twins fell in. So the other twin went inside and told her mom. The moms name was Marie and the twins names were July and Jamie. Jamie was the one who fell in the pond. Marie started crying. She said we had better hide before she comes to get one of us. They did hid byt Jamie found them. Jamie took July and pushed her in the pond. So Marie was alone. The next day the twins came back and they pushed the mother in the pond. And all three of them came back and haunted the house.

The End

There were other signs of growth in the areas of text content, text organization and language. The content moved from less dependence on the picture text to using the picture as a setting for a narrative. The text organization changed from "patterned beginnings," with "disjointed" organization to "well developed" narratives with an emphasis on the story genre. Second graders relied more on the informational genre in their writing, while fourth graders used the story genre more frequently. The language changed from "stilted" language to "firmer" tone with a "sense of authority."

Summary

The success of the raters to establish those texts written in the first and second halves of the study demonstrated that growth in literacy can be established within grade levels without providing prototypes of competency for readers. It also demonstrated that if children are provided the opportunity to write frequently in a supportive environment, growth in writing ability is evident, even to untrained raters. The writing profiles of the students demonstrated that grwoth and development of writing ability is not the same for all children, nor is the process necessarily one of steady progression.

The fact that these students did not receive any guidance from the teachers before or after the writing event demonstrates how powerful the act of writing is and provides evidence that growth

and development of literacy requires engagement in the process. Providing children the opportunity to write may be more important than providing instruction in how to write.

## REFERENCES

Graves, D.H. (1983) Writing: Teachers and children at work. Exeter, New Hampshire: Heinemann Educational Books.

Myers, M.(1980) A procedure for writing assessment and holistic scoring. Urbana, IL: National Council of Teachers of English.